Antibody Aggregation: Insights from Sequence and Structure

Wei Li [1],*, Ponraj Prabakaran [2], Weizao Chen [1], Zhongyu Zhu [1], Yang Feng [1] and Dimiter S. Dimitrov [1]

[1] Protein Interactions Section, Cancer and Inflammation Program, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 21702, USA; chenw3@mail.nih.gov (W.C.); zhuzhongyu@mail.nih.gov (Z.Z.); fengya@mail.nih.gov (Y.F.); dimitrdi@mail.nih.gov (D.S.D.)

[2] Intrexon Corporation, Germantown, MD 20876, USA; praba.ponraj@gmail.com

* Correspondence: wei.li3@nih.gov; Tel.: +1-301-846-6275

Abstract: Monoclonal antibodies (mAbs) are the fastest-growing biological therapeutics with important applications ranging from cancers, autoimmunity diseases and metabolic disorders to emerging infectious diseases. Aggregation of mAbs continues to be a major problem in their developability. Antibody aggregation could be triggered by partial unfolding of its domains, leading to monomer-monomer association followed by nucleation and growth. Although the aggregation propensities of antibodies and antibody-based proteins can be affected by the external experimental conditions, they are strongly dependent on the intrinsic antibody properties as determined by their sequences and structures. In this review, we describe how the unfolding and aggregation susceptibilities of IgG could be related to their cognate sequences and structures. The impact of antibody domain structures on thermostability and aggregation propensities, and effective strategies to reduce aggregation are discussed. Finally, the aggregation of antibody-drug conjugates (ADCs) as related to their sequence/structure, linker payload, conjugation chemistry and drug-antibody ratio (DAR) is reviewed.

Keywords: monoclonal antibodies; antibody aggregation; protein unfolding; antibody domains; antibody drug conjugates

1. Introduction

The tertiary structure of a protein with a given amino acid sequence is defined by competing molecular-scale interactions, which balance the contributions of fold-favoring interactions, such as electrostatic attraction, hydrophobic interaction and hydrogen bonding, to those of unfolding-favoring interactions, like the geometric constraints of chemical bonds, the avoidance of steric clash and electrostatic repulsion [1]. Under native conditions, proteins fold quickly (in a time frame of μs to ms [2]) from the unfolded state (U) to the folded structure (F) in a cooperative manner with several short-lived, meta-stable intermediates (molten globule states) in the down-hill free energy landscape [3]. Proteins experience constant sampling between the folded and partially-folded structures [4]. Under stress wherein either the folded state (F) destabilizes (G_F increasing and $\Delta G_{U\text{-}F}$ decreasing) or the (partially) unfolded structure becomes stabilized (G_U decreasing) [5], proteins could populate partially unfolded conformations and result in aggregation. Protein aggregation is a process in which protein molecules self-associate with each other. Non-native aggregation forms via the strong non-covalent contacting of protein molecules and is thermodynamically stable and irreversible [6]. Aggregation requires proteins to experience unfolding or partial unfolding to present key stretches of residues (so-called "hot spots") to achieve strong interactions between monomers, which often

constitutes a rate limiting step before nucleation growth for aggregation [7]. In many cases, protein aggregates use the energetically-favorable β sheets as the building modules [8]. A common example for this kind of aggregation is amyloidosis, in which the polypeptides build fibril aggregates by stacking against β sheets along the fibril axis [9].

mAbs-based pharmaceuticals have enjoyed increasing success in therapeutic markets [10] and typically target high impact areas, such as cancers, infectious diseases, auto-immune diseases and metabolic disorders [11]. mAbs bear many therapeutic metrics, such as high binding affinity and specificity, long circulation half-life in blood stream, non-toxic nature and easy manufacturing. Although highly desired, one bottleneck limiting mAbs therapeutics' development is aggregation [12,13]. mAbs with 12 sub-domains, large hydrodynamic radii and surface areas, non-symmetrical hydrophobicity and charge distributions are prone to aggregation [14,15]. The immunoglobulin Greek-key β sandwich folding of mAbs is susceptible to edge-edge association [16]. Besides, complementarity determining regions (CDRs) of mAb responsible for antigen binding can also contribute to aggregation due to the frequent occurrences of hydrophobic and electrostatic residues [17,18]. Furthermore, the extensive hydrophobic patches on the surfaces of mAbs, especially on Fc could mediate aggregation [19,20]. These aggregation propensities are amplified by the natural bivalency of mAb. Importantly, the aggregation of mAb could be increased when administered by subcutaneous (SC) delivery in a high mAb concentration of >100 mg/mL [21]. At such high concentrations, mAbs are more susceptible to aggregation [22]. Furthermore, antibody aggregation is highly undesirable, because it could compromise biological functions [12], induce immune responses by breaking B-cell tolerance [23,24] and evoke antibody clearance machinery in vivo [25]. These disadvantages make the control of antibody aggregation imperative in the route to developing successful therapeutics.

Although the mechanism underlying antibody aggregation is generally not fully understood, the aggregation propensity for a given mAb is a function of solution conditions, such as temperature, pressure, pH, ionic strength and excipients (osmolytes, surfactants). Formulation optimization is commonly used to reduce aggregation [26,27]. On the other hand, the susceptibility to aggregation is pre-defined by the intrinsic properties of the antibody, including primary sequence and tertiary structure. Protein aggregation needs some degree of conformational distortion or partial unfolding of the native monomer to expose the aggregation-prone residues to form strong inter-molecular interactions [28]. Antibodies usually contain more than one aggregation-prone region (APR) [29]. Therefore, by contrast to some simple proteins, in which aggregation occurs in a high cooperativity manner [30], antibody aggregation usually occurs through several intermediate states [31], indicating that multi-domain antibody unfolding and aggregation could be understood by analyzing the aggregation of individual domains. The studies of aggregation mechanisms and resistant strategies for antibody domains have inspired the related research of full-length antibodies, although the results obtained from the antibody domains could not always be transferred to the full-sized antibodies. In this paper, we review the impact of sequences and structures on the aggregation of both full length antibodies and antibody domains while discussing aggregation resistance strategies through rational designs.

Antibody-drug conjugates (ADCs) are an important class of therapeutics in oncology. Two ADCs have been recently approved by the U.S. Food and Drug Administration (FDA), including Kadcyla (ado-trastuzumab emtansine) for the treatment of breast cancer and ADCETRIS (brentuximab vedotin) for combating relapsed Hodgkin's lymphoma. ADC is produced by conjugating a toxic reagent to mAb by chemical reactions, which often destabilizes the conformations of mAbs due to the decreased interchain disulfide bonds and the exposed hydrophobic patches. In addition, the hydrophobicity of the linker payloads could expand the APRs, facilitating the aggregation of ADCs. Thus, ADCs are believed to be more prone to aggregation than the parent mAbs. The elucidation of the aggregation behaviors and mechanisms of ADC could help to conquer the aggregation hurdles in ADC development. In this review, we also discuss the aggregation of ADC by focusing on the impact of the drug-to-antibody ratio (DAR) on aggregation.

2. Protein Aggregation

2.1. Why Does a Protein Aggregate?

The native protein has the lowest free energy and the most stable conformation [3] (Figure 1). This is maintained by a spectrum of fine interactions: hydrogen bonds of the main chains and side chains; van der Waals and hydrophobic interactions; the constraints of energetically-unfavorable bond torsions and steric clashes; maximizing chain entropy; electrostatic repulsion and attraction; and interactions between amino acids and the solvent [32]. These interactions collectively balance the protein conformations [33]. Under conditions in which the folded states overwhelm the unfolded states, protein monomers populate the native conformations. In such a case, the nature of the energy barriers for the transient state (TS) prevents protein ensembles from populating aggregation-prone states, and thus, the protein remains soluble without a risk of misfolding and aggregation [34]. However, these energy barriers are compromised under stressing conditions, wherein the decreased G_{TS} combines with the increased G_F kinetically and thermodynamically shifts the "on-way folding" pathway to the "off-way aggregation" pathway, allowing proteins to experience an irreversible aggregation route [5].
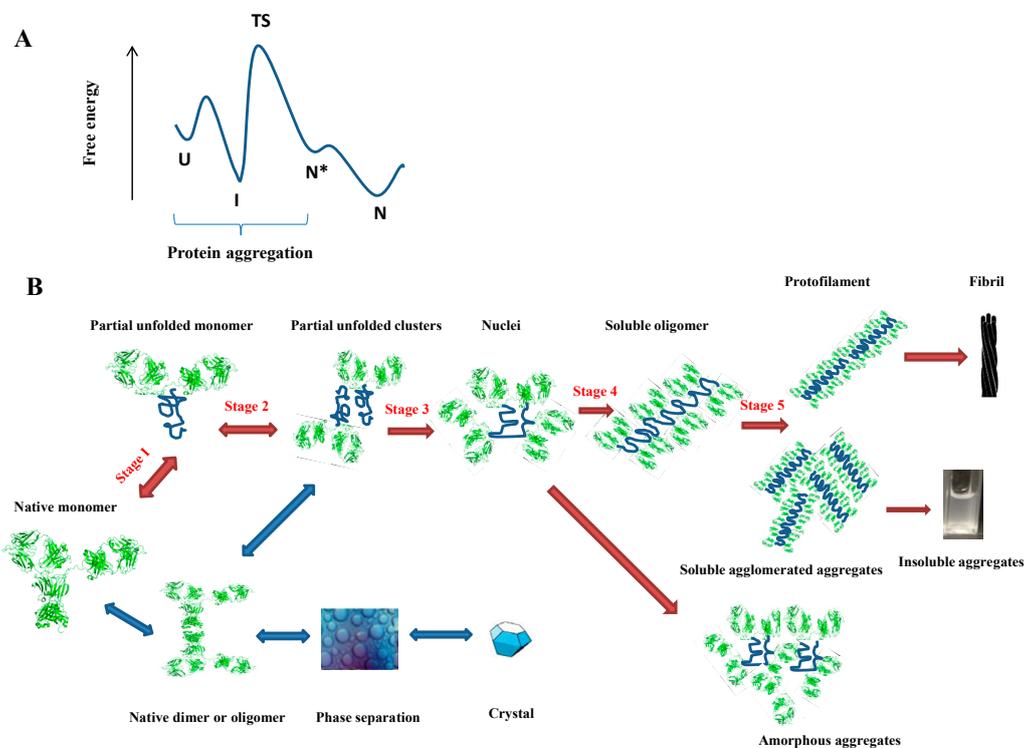
Figure 1. The process of protein folding and aggregation. (**A**) Folding is presented based on the classical thermodynamic and kinetic principles. U, I, N*, N, TS refer to the unfolded state, partially-folded state, locally unfolded state, folded state and transient state, respectively. The conformational ensembles of U, I, N* are prone to aggregation; (**B**) Schematic representation of the protein aggregation process and the possible involved intermediates. This figure uses a monoclonal antibody as an example, but the general behaviors and principles are also applied to other proteins. The red arrows represent the non-native aggregation, while the dark blue arrows denote the native aggregation. The bidirectional arrows show the reversible steps, and the mono-directional arrows account for the irreversible process.

2.2. How Does a Protein Aggregate?

In general, the aggregation process mediated by folding intermediates could be divided into the following five stages (Figure 1B): (I) partial unfolding of the native monomer; (II) reversible self-association of the partially-unfolded or folded protein; (III) net irreversible aggregation nuclei

formation; (IV) further aggregation growth by monomer chain addition; (V) aggregation association to form the high molecular weight soluble aggregates or insoluble precipitation [35].

Protein aggregation mediated by folding intermediates is triggered by the unfolding of native monomer. The unfolding energy barrier ($\Delta G_{U\text{-}F}$) that allows the samplings of partially-disordered conformations needs to be overcome in Stage I of aggregation [28]. $\Delta G_{U\text{-}F}$ not only intrinsically correlates the protein associated G_U and G_F, but also depends on temperature (T), pressure (*p*) and the conditions of the solution [36]. Stage II involves the association of unfolded monomers, which is governed by the protein colloidal interactions implicated by the hydrophobicity and charge distributions of proteins [37,38]. The key stage for protein aggregation lies in Stage III, the nucleation step, in which the unfolded protein experiences the structural re-arrangement, such as the alteration of surface charge distributions, the exposure of the aggregation-prone regions (mainly referring to the hydrophobic patches to the solvent) and changes of the topologies for the β sheet regions or even re-orientation of the α-helix into the β-strand [5]. This step usually constitutes the rate-limiting step in the protein aggregation, after which the subsequent growth stages are much faster [7]. The existence of this nucleation stage is rationalized by the experimentally-observed "lag time" during protein aggregation, which could be shortened by "seeding" the preexisting aggregates [39]. The overall protein aggregation occurs kinetically rather than dynamically and, therefore, is pathway dependent [7]. Thus, it is often experimentally observed that proteins bear good thermostability (low G_F), but indeed aggregates readily due to kinetically-favorable conditions [40].

2.3. How to Mitigate Protein Aggregation?

Although protein aggregation is a priori unpredictable kinetically, the thermodynamic aggregation potentials could be inferred from the intrinsic characteristics of protein sequences and structural features [7,18]. Therefore, aggregation could be well controlled by the rational design of protein sequences and structures. Protein unfolding and the nucleation are usually the key steps for protein aggregation, and the subsequent monomer addition and aggregate association occur much faster. Thus, strategies ameliorating protein aggregation often target the aggregation Stages I, II and III. The corresponding methods for mitigating aggregation thus are: (1) stabilizing the native monomer (decreasing G_F) or destabilizing the partially-unfolded monomer (increasing G_U) to reduce the potential of protein unfolding at Stage I; (2) altering the protein surface charge distributions to increase the electronic repulsion between the unfolded monomers at Stage II; and (3) disturbing the structural re-arrangements of unfolded monomers in Stage III to disfavor hydrophobic contacts and the packing of β strands. These strategies could be learned from the nature of protein structures and mechanisms. For example, many proteins involved in hereditary forms of protein deposition diseases bear mutations decreasing the conformational stability of the folded monomer and promote aggregation in vitro [38,41,42]. In addition, the native proteins disfavor the sequences of alternating polar and non-polar residue for β strand assembly [43], as well as clusters of many consecutive hydrophobic residues to decease the tendency of aggregation before folding [44]. Antibodies either inwardly point a charged residue in the middle of the β strand to disfavor the hydrophobic associations or they locate a proline to introduce a bulge dissecting the strand to avoid edge-edge association [16].

2.4. Computational Methods for Studying or Predicting Protein Aggregation

Protein aggregation involves various aggregate intermediates and pathways. Many computational models have emerged to dissect the aggregation mechanism and to evaluate how the external factors, pH, ionic strength, etc., influence the aggregate intermediates. In this regard, many groups exploited the coarse-grained (CG) lattice models to study the aggregation mechanism, in which protein molecules are treated as on-lattice single chains, and residues represented are as beads [45,46]. Their interactions are calculated in specifically pre-defined force fields, and the molecular association equilibrium is simulated using conformation searching algorithms, such as Monte Carlo (MC). To obtain more kinetical details, many groups have performed the more accurate atomistic simulations by simplifying proteins into the peptides bearing high aggregation propensities. These molecular simulations provide

insights into exploring the oligomeric conformations that can seed the aggregation [47] and figuring out how peptide side chains kinetically and dynamically affect the amyloid aggregation [48].

On the other hand, many computational algorithms have been developed to predict APRs, which usually have unique sequences regarding charge, aromaticity, hydrophobicity and secondary structural propensity [49,50]. These calculation tools could be classified into two types. One is the statistical algorithms to rank the propensity of aggregation for stretches of amino acids by comparing them to the existing polypeptide databases composed of amyloidogenic peptides, which includes Aggrescan (Aggregation Scan), PAGE (Prediction of Aggregation), TANGO (a statistical mechanics algorithm), Zyggregator, Amylpred (Amyloid Prediction), etc. [18,51,52]. Most of these tools only use the protein sequence as input to find out the short APRs with 5–9 residues prone to forming amyloid-like fibrils. Although these methods have achieved some success, one should be cautious about the APRs identified by these methods, since these APRs are just necessary, but not sufficient for forming aggregates. To facilitate aggregation, APRs need to be exposed to contacting the neighboring molecules. Therefore, APRs prediction should be verified by the experimental results when the APRs do not exist in the solvent exposed loops or on the surface of folded proteins. The other method for identifying APRs is the molecular simulation (MD), which ranks protein conformations according to their aggregation propensity. Cecchini et al. have used MD to predict APRs in human amyloid β-peptide, amylin and the prion protein PrP Ure2p$_{1\text{-}94}$ [52]. Recently, Chennamsetty et al. have developed a full atomistic MD simulation method (spatial aggregation propensity (SAP)) to identify aggregation-prone motifs with surface exposed hydrophobic residues on full IgG1 [53]. By using SAP, they have successfully achieved mitigating the aggregation of IgG1 by mutating those APRs [54]. Collectively, these calculation methods could provide convenient methods for ranking protein candidates in the early development process, which could guide the design of aggregation-resistant proteins.

3. Antibody Aggregation

### 3.1. Sequences and Structures of mAbs

The overall architecture of a typical IgG consists of two identical light chains and heavy chains. Each light chain folds into two domains, V_L (variable light) and C_L (constant light), while each heavy chain contains four domains of V_H (variable heavy), C_H1, C_H2 and C_H3 (Figure 2) [55]. The whole IgG forms three structural units of equivalent size, two Fabs and a Fc dimer. Each Fab is composed of V_L, C_L, V_H and C_H1, and Fc is a homodimer of C_H2 and C_H3. Fab and Fc are loosely connected by the hinge region and are not considered to interact with each other. These multi-domains enable the IgG to allocate its functions into different domains. Both V_H and V_L in Fab collectively mediate the antigen binding via CDR loops. Fc is involved in effector functions and pharmacokinetics by interacting with receptors.

The V domains of IgG are the smallest entities for antigen binding. The V domain consists of two layers of β sheets connected by a disulfide bond and by a cross-over connection through the CDR1 loops (Figure 3) [56]. The C domain lacks C' and C'' strands, which would otherwise correspond to CDR2 in the V domain. The CDR3 loops in V_H and V_L are established by the V-D-J and V-J rearrangement of the antibody gene respectively and somatic mutations that contribute to high diversities for targeting various antigens. CDRs are the core part for antigen binding, which contain high frequencies of aromatic and hydrophobic residues, such Tyr, Phe, Leu and Ile, as well as residues functioning as hydrogen bonding donors, such as Ser, Thr, Asn and Gln [18]. V_H and V_L associate with each other via hydrophobic interactions involving residues Val37, Leu45 and Trp47 (according to Kabat numbering). C_H1 and C_L hold together by strong hydrophobic packing between residues Val190, Phe174 and Leu143 in C_H1 and Leu135, Phe116, Phe118 and Val133 in C_L, which combined the V_H/V_L association to constitute the stable Fab [31]. Fab connects to Fc via the flexible hinge region, where the upper region contains several disulfide bonds for establishing inter-chain ligation, and the lower region is usually hydrophobic and participates in the Fc receptor binding. The flexibility of the hinge ensures the domain movement and the orientation of Fab and Fc, which could modulate the antigen binding

and effector function. C_H2 is believed to be the least stable domain in IgG due to the lack of direct interactions between the C_H2 dimer, except the weak side chain interactions of glycans [57]. In contrast, the C_H3C_H3 dimer comes into tight contact with each other by the hydrophobic interactions involving residues Tyr438, Phe436, Leu391 and Leu372 and electrostatic interactions, such as salt bridge of Glu357-Lys370 and Asp399-Lys409 [58]. C_H2 associates with C_H3 via the salt bridges of Lys248-Glu380 and Lys338-Glu430. Collectively, IgG forms a well-folded globular structure via extensive intra-domain and inter-domain interactions.
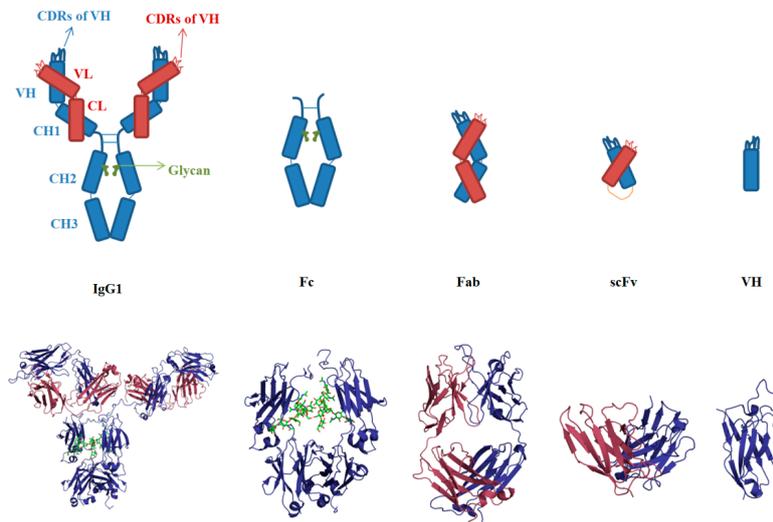
Figure 2. Molecular architecture of an immunoglobulin G1 (IgG1) antibody and its fragments. An IgG consists of two heavy chains (blue) and two light chains (red). The glycan is presented by the green color. Fc (the crystallizable fragment) is a dimer of C_H2, C_H3 and glycans. Fab (antigen-binding fragment) is composed of variable heavy (V_H) and light (V_L) domains, as well as two constant domains (C_H1 and C_L). ScFv is the artificial format containing V_H and V_L connected by a flexible linker (yellow). V_H (or V_L) is the minimal unit for antigen binding mediated by complementarity determining regions (CDRs). The CDR loops in the V_H domain are denoted as H1, H2 and H3 (blue); the CDRs in the V_L domain are named as L1, L2 and L3 (red). Below are the 3D structures of an HIV neutralizing antibody b12 with intact IgG1, Fc, Fab, scFv and V_H formats.
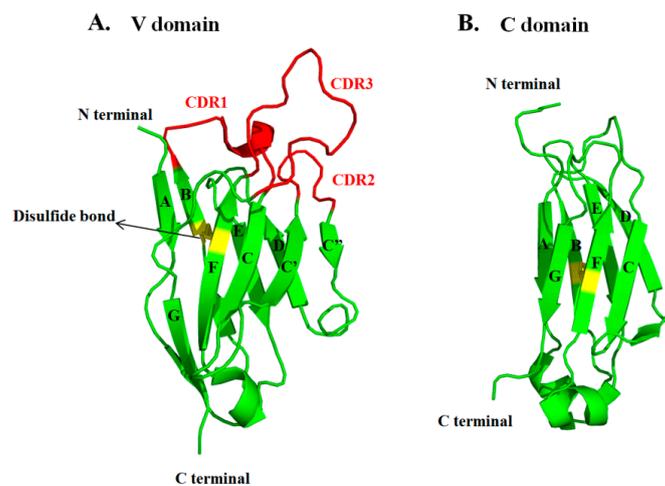
Figure 3. The structural comparisons of the V domain and the C domain. (**A**) and the C domain (**B**). The β sheets are presented by the green cartoon models with the CDRs denoted in red and the disulfide bond in yellow. The anti-parallel β sheets are numbered as A, B, C, C′, C″, D, E, F. The C domain does not contain the C′ and C″ strands.

3.2. Aggregation of Full-Length IgG

IgG contains extensive intra-domain and inter-domain hydrophobic interactions. When subjected to structural fluctuations, those hydrophobic interactions are readily exposed to constitute aggregation nuclei [59]. The subdomains of IgG belong to the immunoglobulin superfamily (IgSF) with β-strand sandwich folding, which is intrinsically prone to amyloid aggregation by edge-edge association [18]. For example, trastuzumab (Herceptin) contains two closing β-strands in the Fab, wherein one strand (SVFIFP) at the edge of the four-stranded β-sheet of C_L is packed against the four-stranded β-sheet of the C_H1 domain, thus mediating aggregation [18]. The intermolecular beta sheet associations were frequently found in the IgG1 aggregate induced by various stress conditions, such as heating and stirring [60]. Besides those non-covalent associations, the free sulfhydryl group derived from the impaired disulfide bonds of IgG promotes aggregation through intramolecular scrambling and/or intermolecular crosslinking [61]. In this regard, particular attention needs to be paid to the unpaired non-canonical Cys for antagonizing aggregation. Buchanan et al. have achieved decreasing the aggregation propensity of Ang2 mAb by mutating a free Cys residue in the proximity of light chain CDR2 (LCDR2) (Cys49) into Thr or Asn [62]. On the other hand, the aggregation susceptibility of IgG is deeply modulated by the features of CDRs. Given that CDRs are also responsible for antigen binding, it is challenging to engineer the CDRs to resist aggregation without compromising antigen binding. To reconcile the antigen binding, aggregation-resistant engineering usually does not directly target CDR bearing the APRs, but rather the edge residues flanking CDRs or the surrounding regions. Wu et al. have improved the solubility of an anti-IL-13 monoclonal antibody CNTO607 by mutating a set of hydrophobic residues (Phe-His-Trp) in heavy chain CDR3 (HCDR3) to Ala. However, the binding affinity of the mutated mAb was decreased significantly (>1000-fold) [17]. Alternatively, instead of engineering HCDR3, the authors have resorted to introducing a hydrophilic glycan into the neighboring HCDR2, which could shield the hydrophobic triad in HCDR3, but meanwhile did not intervene in the antigen binding [17]. Another elegant study came from Dudgeon et al. showing that the incorporation of negatively-charged residues, such as Asp and Glu, into the HCDR1 and LCDR2 did not impact the antigen binding and function of full-length trastuzumab against HER2, probably because it is HCDR3 and LCDR3, rather than HCDR1 and LCDR2, that contribute to HER2 binding [63].

Another factor complicating IgG aggregation is the glycosylation. Usually, the glycans attached at N297 are believed to benefit the aggregation resistance for IgG by shielding hydrophobic residues from being exposing to the solvent. Trout et al. have used SAP calculations to demonstrate that glycan attached at IgG1 Asn297 shields the couples of hydrophobic residues, such as F241 and F243, from exposure to the solvent. Consequently, the aglycosylated mAbs are less stable and therefore aggregate more easily than the glycosylated mAbs [64]. In addition, it is believed that glycoforms could adjust the conformation of Fc in either "open" or "closed" states and, thus, modulate the colloidal interactions between IgG [65]. Schaefer et al. have reported that IgG bearing high mannose derived from yeast is more resistant to aggregation than the counterpart expressed by mammalian cells with complex type glycans [66]. Hence, one can change aggregation liability by altering the glycoform of IgG or introducing additional glycans on IgG [67]. By contrast, the hydrophobic patches in Fc, which constitute the docking sites for other molecules to implement important biological functions, are "hot spots" for aggregation [68]. Trout et al. have exploited SAP to identify the exposed hydrophobic residues spreading across the whole sequences of Fc, such as Leu309 in C_H2 and Ile253 in the C_H2–C_H3 junction [53]. Those hydrophobic patches have been reversed by introducing the mutations of L234K, L235K, I253K and L309K to decrease IgG–IgG self-association and aggregation [54]. Although the different IgG isotypes bear relatively conserved constant fragment and relatively invariant hydrophobic patches, they have intrinsically distinct thermostabilities and colloidal stabilities due to the minor differences of the Fc sequences, the length of the hinge linker, the number of disulfide bonds in hinge regions and the pattern of glycosylation [69,70]. In many cases, it is believed that IgG1 has the highest stability compared to IgG2, IgG3 and IgG4. IgG2 has two more cysteines in the hinge compared to

IgG1 and is prone to the presentation of free sulfhydryl radicals for exacerbating aggregation. IgG3 has a relatively long hinge region, which is susceptible to protease cleavage and renders IgG3 prone to chemical degradations. IgG4 is prone to forming into the bi-specific dimer by domain swapping [71]. Thus, most therapeutic mAbs have adopted the IgG1 forms. However, in special cases, the aggregation propensity of IgG1 could be improved by the isotype switching [67,72].

On the other hand, the domain-domain associations need to be taken into account when studying the aggregation of IgG, which contains substantial molecular-scale interactions in the V_H/V_L, C_H1/C_L, C_H2/C_H3 and C_H3/C_H3 interfaces. The domain-domain interactions confer IgG interface free energy for native folding, which would disappear when one of the interacting domain unfolds [57]. Thus, the unfolding and aggregation of full length IgG is often triggered by the least stable domain. It is believed that C_H2 unfolds first and triggers the aggregation process, while C_H3 is the most stable domain [73]. Fab usually bears the middle thermostability, but is significantly modulated by the CDR sequences [74]. IgG aggregation is most frequently studied at an acidic pH since IgG needs to be exposed to acid for resin purification or virus clearance [35]. The low pH renders the charged residues in the domain interfaces experiencing protonation, which could disturb the hydrogen bonds and salt bridges [59]. Consequently, the polar residues would carry excess charges and destabilize IgG by intramolecular charge–charge repulsions [7]. The low pH was often combined with temperature ramping for evaluating enthalpy changes and the unfolding intermediates during IgG unfolding determined by DSC measurement [57]. The DSC thermograms of IgG usually contain more than one endotherm peak corresponding to the unfolding of its separated domains [73,75], which indicates that the thermostability of the individual domain collectively contributes to the overall aggregation propensity of full-length IgG. Therefore, to explore the strategies for mitigating IgG aggregation, one could scrutinize the factors influencing aggregation on the basis of its separated domains. Studying the aggregation of antibody domains is more straightforward and informative. The aggregation-resistant lessons learned from single or multi-antibody domains could guide the aggregation-resistant engineering of full-length IgG. Therefore, studying the aggregation of antibody domains or fragments has significant relevance for combating the aggregation of full-length IgG. In the next section, we will discuss the relationships between aggregation propensity and the structures of antibody domains, as well as their aggregation-resistant strategies.

3.3. V_H and V_L

V_H and V_L as single domains are the smallest units for antigen binding. These single domains have attractive advantages as diagnostic and therapeutic reagents due to their small size, which confer them capacity to access cryptic epitopes and enhance penetration into solid tissues [76]. V_H and V_L are intrinsically prone to aggregation due to the exposure of the hydrophobic V_H–V_L interface. V_H has been shown to be more prone to aggregation than V_L, and its aggregation is more dependent on its CDRs than V_L [77]. The stabilizing free energy for V_H and V_L results from the hydrophobic core inside the immunoglobulin β barrel and the disulfide bond connecting the two β sheets layers. While an intra-domain disulfide in the antibody domain contributes a range of 4–6 kcal/mol of free energy to the folding [78], 1 Å^2 of the hydrophobic contact corresponds to 25 cal/mol [79]. The CDRs responsible for antigen binding also impact the colloidal stability and aggregation of V_H and V_L. Generally, the aggregation resistance strategies could be divided into rational and evolutionary approaches. While rational approaches exploit structure-based knowledge or sequence information to guide the aggregation-resistant mutations, the evolutionary methods involve the selection of a single domain antibody library by stability-improving pressures, such as temperature and pH. These two approaches could be combined to design aggregation-resistant single domain antibodies (sdAb) (Figure 4).

The elucidation of structural and sequence determinants underlying aggregation propensity diversifies the rational approaches of combatting the aggregation of sdAb, which include the engineering of the overall charge, CDRs, disulfide bonds, V_H–V_L interface and the residues in framework (FR) regions influencing aggregation. Antibodies with net charges are less prone to aggregation due to the intermolecular repulsions. Tanha et al. have reported that the

aggregation-resistant human V_Hs obtained by panning against target enzymes using a transient heat denaturation approach bear acidic pIs, similar to the naturally-occurring camelid V_HH and shark V_{NAR}, which is reminiscent of protein acidification constituting a universal mechanism to confer functional sdAbs [80]. The same principle may be also applied to the full-length IgG. Schaefer et al. have shown that IgG expressed in yeast bearing the "EAEA" sequence at the N termini of the light chain and heavy chain are more aggregation resistant than the counterpart expressed in mammalian cells [66]. Consistent with this, introducing the negatively-charged residues into CDRs is a feasible route for mitigating the aggregation of V_H and V_L. Perchiacca et al. have improved the solubility of V_H containing aggregation-prone Aβ peptides by inserting the "DED" triad into the edge of CDR3 [81]. Dudgeon et al. have found that incorporations of negatively-charged residues, especially Asp, into the CDR1 of V_H (Positions 28, 30–33, 35) and CDR2 of V_L (49, 50–53, 56), are universal strategies for conferring aggregation resistance to sdAbs [63]. In addition to the introduction of negatively-charged residues, CDRs are subjected to extensive aggregation-resistant engineering since CDRs are hydrophobic "hot spots" mediating the aggregation of sdAbs. Rational designs usually include the mutagenesis of hydrophobic residues into hydrophilic ones [17,82] and the introduction of cysteines into CDR3 to constrain the conformation of the long protruding CDR3 or to mediate cross-linking with CDR1 to stabilize sdAbs [80,83,84]. On the other hand, the exogenous disulfide bond could be introduced into the FR regions to strengthen the thermostability of V_H and V_L. Kim et al. have reported that the introduction of the non-canonical disulfide bond between Cys54 and Cys78 residues increased the thermostability of V_H by 14–18 °C [85]. In addition, the same authors have also achieved improving the thermostability of V_L by 5.5–17.5 °C through the creation of disulfide bonds linking residues 48 and 64 [86]. Another aggregation-resistant method involves the engineering of the exposed V_H–V_L interface that would otherwise be buried in intact antibodies or other bigger fragments. Indeed, hydrophilic mutations near the former V_H/V_L interface have been demonstrated to improve the solubility of dAbs [87]. The naturally-occurring V_HH from camels has inspired the screening of aggregation-resistant mutants of V_H, although the camelized mutations are often not applied to therapeutic mAbs considering the potential immunogenicity [88]. Furthermore, accumulating research has established the impact of the residues in FR regions on the aggregation of sdAbs. For example, the residues of Glu or Gln6 [89], Arg66 and Gln105 [90] in V_H and residues of R24, Y49 [63] and Pro8 [91] in V_L are the key players in mediating the aggregation of sdAbs.
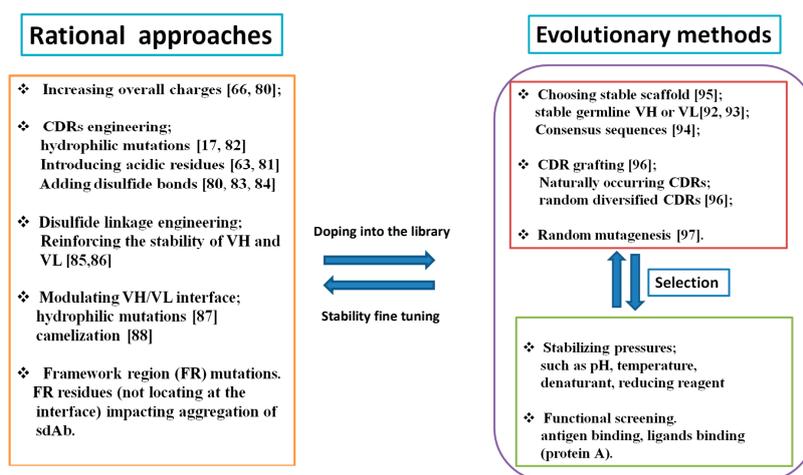
Figure 4. Rational approaches and evolutionary methods for decreasing the aggregation of human single-domain antibodies. These two methods could be combined to develop aggregation-resistant V_H or V_L. The rationally-designed mutations can be doped into the library to increase the possibility of identifying binders with superior biophysical properties. The binders obtained from such a library could be subjected to stability fine-tuning and further guided by rational approaches.

Alternatively, the library-based evolutionary approaches could be used for alleviating the aggregation of sdAbs. Antibody domains of different sub-families bear distinct thermostability and aggregation propensity. The V_H1, 3 and 5 gene families generally displayed better physical characteristics than the 2, 4 and 6 gene families [92]. For V_L, V_κ is more desirable than the V_λ chain, with V_κ3 being the most thermodynamically-stable followed by the V_κ1 and V_κ4 [93]. Besides, based on the sequence and structural comparisons, some groups have found the so-called "consensus sequences" benefiting the folding of V_H, which usually bear high frequencies in the repertoire [94]. By combing the favorable germline domains and those consensus sequences, one could establish stable scaffolds to construct libraries with randomized CDRs, which could be further subjected to some pressures to select the biophysically-desired binders [95]. Our group succeeded in isolating a stable V_H antibody, m36, for targeting HIV-1 CoRbs by panning a V_H phage library, which is composed of a stable scaffold based on the V_H3 with the incorporation of partially-randomized CDR1 (Positions 30 and 31) and naturally-occurring CDR2 and CDR3 [96]. On the other hand, one could perform random mutations for the selection of stability-improved sdAbs, in the case where the parent sdAb for a given antigen has already been developed [97].

3.4. scFv

scFv consists of the variable domains of V_H and V_L connected by a flexible linker, which combines with the strong inter-domain associations to ensure that scFv retains the antigen binding of the full-length IgG if no reorientation of the V_H and V_L occurs. Due to its small size (~30 kD), scFv bears many therapeutic merits as a diagnostic and therapeutic agent, such as easy expression, relative stability compared to V_H and deceased toxicity when used in radioisotopes and bio-imaging [98]. Like the separated V_H and V_L, the thermostability and aggregation liability of scFv relates tightly to the CDRs' sequences and the intrinsic stabilities of the sub-domains [99]. Thus, the general aggregation-resistant strategies for V_H and V_L are also applied to the scFv. General strategies include introducing the stabilization mutations, such as residues 6 and 66 in V_H, Pro8 in V_κ [89,91], choosing stable frameworks of V_H and V_L, such as the combinations of H3$_\kappa$3, H1b$_\kappa$3, H5$_\kappa$3 and H3$_\kappa$1 for constructing scFv [92] and introducing CDR mutations (see the above *V_H and V_L* Section). However, as a fragment with strong V_H/V_L associations, scFv has its unique features. The V_H/V_L interface of scFv is usually conformationally dynamic and experiences "transient opening", which could expose the hydrophobic patches to promote aggregation [100]. Therefore, stabilizing the V_H–V_L interface has enjoyed many successes for improving the solubility of scFv. Corresponding strategies include introducing a disulfide bond into the contact interface [101], forming new salt bridges and hydrogen bonds between V_H–V_L [102] and adding "knob-into-hole" mutations [103]. Another influential factor impacting scFv aggregation is the monomer-dimer-multimer equilibrium occurring via the domain-swapping manner, where the V_H domain of one scFv pairs with the V_L domain of another scFv and vice versa [104]. This domain swapping highly depends on the linker length in which the short linker of 5–10 residues forms a stable dimer (diabody), and the long linker of 15–20 amino acids favors a monomer [99]. Thus, optimizing the linkers between V_H and V_L by changing both the linker sequences and length could alleviate the aggregation of scFv [105]. Including linkers of 25 amino acids is reported to be a common way of increasing the stability of scFv by populating monomer species [99]. Besides, the domain swapping is also impacted by the expression condition, solution ionic strength and pH [106].

3.5. Fab

Fab contains four domains, namely, V_H, C_H1, V_L and C_L, in which V_H/V_L and C_H1/C_L strongly associate via hydrophobic contact and hydrogen bonding, while few interactions exist at the V_H/C_H1 and V_L/C_L interface. Fab has unique advantages as the diagnostic tool and pharmaceutic reagent due to its small size, easy production and relatively longer in vivo half-life than the single domain antibody [107]. Fab-based therapeutics is currently on the market, e.g., ReoPro (Centocor), Lucentis (Genetech) and Cimzia (UCB) [108]. Generally, Fab is more stable than V_H and scFv since the

hydrophobic interfaces exposed in V_H and scFv are buried by the constant domain in Fab. Like sdAbs and scFv, the unfolding and aggregation propensity of Fab are also impacted by CDRs. It is reported that IgG pools from human blood exist as dimers due to the association of the distal ends of their Fab arms, similar to the idiotype-anti-idiotype complexes [109]. The properties of CDR residues could modulate the colloidal stability, thus mediating the self-association of Fab in the native condition [110]. Besides CDRs, the sequences and structural features of FR regions also impact colloidal interactions by altering the charge distributions across the Fab [20]. Meanwhile, colloidal interaction could be deeply modulated by the solution pH and ionic strength [108]. Another factor complicating Fab aggregation is the inter-domain disulfide bond. The disulfide bond at the C terminus of C_H1 and C_L of IgG1 can lock domain conformations and strengthen the inter-domain interactions. By mimicking this disulfide bond, Peters et al. have improved the thermostability of Fab and the disulfide bond heterogeneity of full-length IgG4. This disulfide bond was formed through mutating C127 in the N terminus of C_H1 to a serine and simultaneously introducing a cysteine at the C terminus of C_H1 (Positions 227–230), which facilitates disulfide bonding to the C_L [111].

On the other hand, unlike sdAbs, Fab structurally features two inter-domain interactions, which are targets for aggregation-resistant strategies. The domain-domain cooperativity of V_H/V_L and C_H1/C_L in Fab is highly predisposed in the unfolding process, as shown by the single and sharp endotherm peaks in the DSC thermograms [57]. The C_H1/C_L interface is believed to be more stable than the V_H/V_L interface, but also depends on the characteristics of CDRs. While the V_H/V_L interface is pre-defined for a given Fab, which has limited spaces for engineering considering the requirement of attaining the antigen binding, the C_H1/C_L interface constitutes a convenient platform for improving inter-domain interactions. Teerinen et al. have reported that the solvated hydrophobicity of C_H1/C_L could be increased by mutating Thr178 to Val or Leu, which leads to increased unfolding free energy of Fab [112]. Recently, our group has successfully used an evolutionary method to introduce a set of mutations in the C_H1/C_L interface including S64E, S66V in C_H1 and S69L, T71S in C_κ. These mutations conferred C_H1/C_L with improved inter-domain associations due to enhanced hydrophobic contact and hydrogen bonds [113]. The improved variant could retain its homogenous monomeric state even in the presence of 1 mM TCEP, indicating that the lack of disulfide bond has no impact on the non-covalent inter-domain interactions (Figure 5).
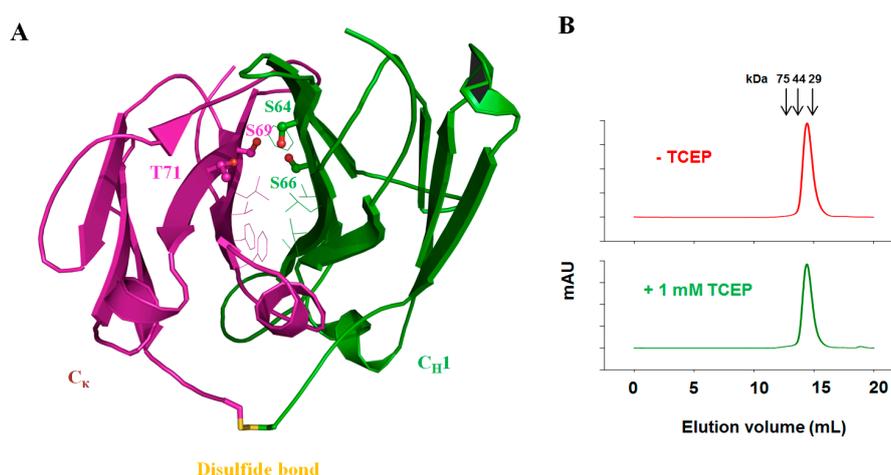
Figure 5. Rational design and identification of stabilized C_H1–C_κ. (**A**) Structural analysis of the C_H1–C_κ interface. The side chains of hydrophobic residues at the interface are shown in slim stick representation. The four amino acid residues lining a void structure are indicated with their side chains shown in a bold ball-and-stick representation; (**B**) Size-exclusion chromatography of mD1.22-C_H1/m36.4-C_L variants. Proteins were treated with and without 1 mM TCEP (tris(2-carboxyethyl)phosphine) before analysis. The arrows at the top indicate the elution volumes of the molecular mass standards in PBS (pH 7.4): carbonic anhydrase (29 kDa), ovalbumin (44 kDa) and conalbumin (75 kDa).

3.6. Fc

The hydrophobic regions contain APRs promoting antibody aggregation. For IgG, more APRs are found in Fc than in Fab, as evidenced by the molecular simulations performed by Bernhardt et al. [53]. By using the so-called SAP method, they have found as many as 14 aggregation-prone motifs in the IgG1 constant domains. These motifs contain one to seven residues and are largely conserved among all IgG subclasses (IgG2, IgG3 and IgG4). According to their research, most of those motifs are located at the lower hinge region and C_H2–C_H3 interface of Fc. Indeed, the hydrophobic lower hinge (sequence: 231-APELLGGPSVFLFPP-245) is not only the "hot" site for engineering IgG effector function, such as antibody dependent cell mediated cytotoxicity (ADCC) and complement dependent cytotoxicity (CDC) [68], but also a focus for improving antibody solubility and thermostability. In Bernhardt et al.'s later study, the authors have improved the solubility of full-length IgG by mutating the hydrophobic residues of L234 and L235 into lysine. This improvement probably results from the enhanced thermostability of C_H2 domain as revealed by the DSC analysis [54]. Gong et al. have achieved improving the thermostability of the C_H2 domain and decrease aggregation propensity by removing the unstructured loop composed of sequence of N-APELLGG-C [114]. Furthermore, the same authors have found that the aggregation of C_H2 could be significantly decreased by mutating APRs residues, such as F241, F243 identified by the TANGO program into hydrophilic residues (personal communications). Another hot spot for combating aggregation lies at the C_H2–C_H3 interface, which bears high adaptability to binding different ligands, such as protein A/G, FcRn and rheumatoid factor [115]. DeLano et al. have shown that the C_H2–C_H3 junction contained a hydrophobic consensus motif composed of Asn434, Ile253, Met423, Tyr436, Met252 and Ser254 for accommodating different natural scaffolds, such as two α-helices of the B-domain of protein A and two β-strands (engineered peptide with sequence of DCAWHLGELVWCT) [115]. By mutating I253 to lysine, the solubility and stability of full-length IgG1 could be improved according to Bernhardt et al.'s study [54]. Another elegant piece of evidence clarifying the implications of this hydrophobic patch in IgG aggregation came from Kolenko et al.'s research, showing that, in the crystal structure of mouse IgG2b-Fc, the C_H2–C_H3 interface interact with the C'E loop (the residues Arg293–Thr299) containing glycosylated Asn297 and another nearby loop of the same neighboring molecule [116]. This complex of two antiparallel-oriented dimers of the Fc provided a structural model of Fc:Fc-mediated immunocomplex (IC) formation and increased aggregation. The involvement of the C'E loop in Fc–Fc association may also have relevant for the roles of glycans in the stabilization and aggregation of IgG. Chen et al. have observed enhanced thermostability and decreased aggregation of full-length IgG1 after introducing an engineered aromatic sequon (EAS) (Q295F/Y296A) into the glycosylated C'E loop [117]. Although the aggregation resistance was ascribed to the improved stability of C_H2 due to the interactions between F295 and GlcNAc1 and core fucose, it may be interesting to explore how the altered C'E loop impacts the Fc–Fc interactions and subsequent colloidal aggregation of IgG1.

Fc aggregation depends on the thermostability of C_H2 and C_H3. Actually, C_H2 is considered to be the least stable domain and usually triggers the unfolding of full-length IgG [118]. Therefore, improving the thermostability of C_H2 and C_H3 could benefit the stability and solubility of the full-sized IgG. One effective strategy for improving the stability of C_H2 and C_H3 is modulating intra-domain disulfide bonds. Gong et al. have increased the transition melting temperature (T_m) of C_H2 from 54.1 °C to 73.8 °C by introducing a disulfide bond in Positions 242 and 334 (m01) [119]. Ying et al. have achieved improving the T_m of a monomeric C_H3 from 41 °C to 76 °C by introducing a disulfide bond connecting Positions 343 and 431 [120]. Wozniak-Knopp et al. have reported that the human IgG1 Fc could be stabilized by introducing intra-domain disulfide bonds in the C_H3 domain. In their research, two engineered disulfides with one linking the N-terminus of the C_H3 domain with the F-strand and the other connecting the BC loop and the D-strand collectively led to an increase of the T_m of ~15 °C for IgG Fc [121]. Collectively, Fc with widely-distributed APRs has gained much attention for aggregation resistance, which benefits the design of full-length IgG with improved biophysical properties.

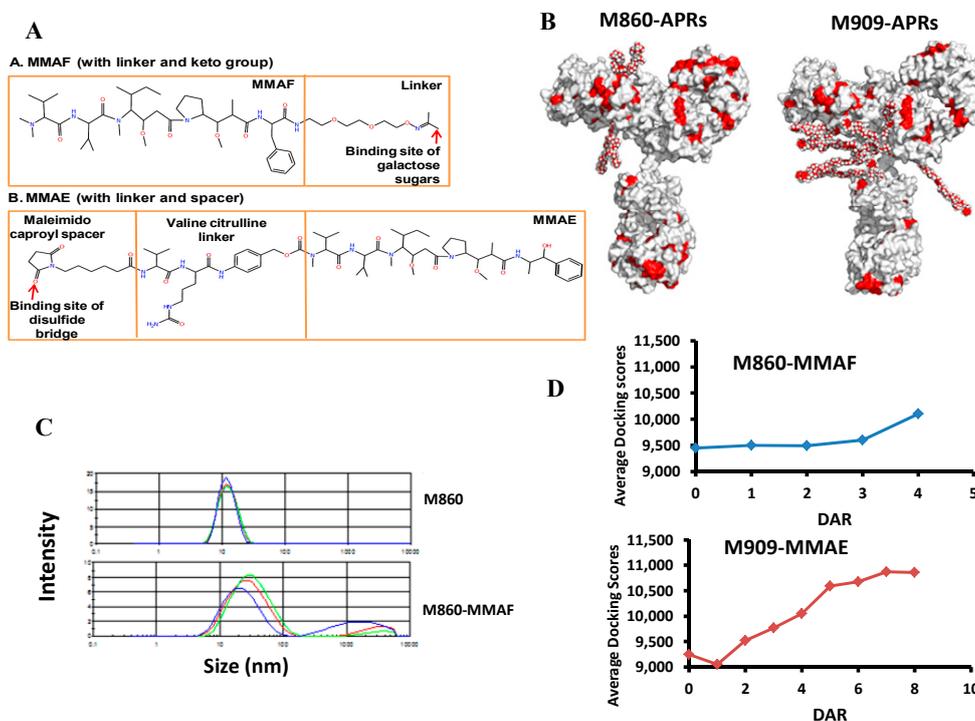## 4. Aggregation of Antibody Drug Conjugates

Antibody drug conjugates (ADCs) combine the specificity of monoclonal antibodies (mAbs) with the potent cytotoxic activity of small molecule drugs for the treatment of cancer and other diseases [122]. The development of ADCs has been significantly hampered due to the strong tendency of ADCs to aggregate or "clump up" [123].

ADCs have three major components, antibody, linker and the cytotoxic drug as the payload. Clearly, the aggregation propensity of ADC is directly related to the parent mAbs, as exemplified by Beckley et al. showing that eight kinds of ADCs with the same linker-payloads of vc-MMAE (monomethyl auristatin E), but with distinct variable domains had different propensities of forming high molecular weight species (HMWS) [124]. Meanwhile, the aggregation of ADC is complicated by the conjugation chemistry and the nature of linker-payloads. The conjugation reaction usually involves the activation of side chains of cysteine, lysine or the keto group of carbohydrate from mAb and subsequent ligation with functionalities from the linkers. During the production process, the activated intermediates containing free sulfhydryl group of Cys and the maleimido group from the linker could mediate the inter-molecular crosslinking and lead to aggregation. Wakankar et al. have demonstrated that a lysine-activated intermediate, Tmab T-MCC (Trastuzumab-maleimidylmethyl cyclohexane-1-carboxylate), was more prone to aggregation than the fully-conjugated ADC T-DM1 (Trastuzumab emtansine) [125]. Additionally, the conjugation type could impact the conformation of mAbs and alter aggregation potentials. It was reported that thiol ligation could induce more aggregation than lysine coupling [126], probably due to the reduction of the inter-chain disulfide bonds in thiol coupling.

For the impact of the linker-payload, their hydrophobically-aliphatic moieties expand the APRs on ADC, facilitating the aggregation by forming hydrophobic patches. Guo et al. have pointed out the contributions of linker payload to the overall hydrophobicity of the ADC by molecular modeling [127]. The increased hydrophobicity of ADC mediated by the linker payload is also supported by the hydrophobicity calculations and APR predictions of two ADCs developed in our group, m860-monomethyl auristatin F (MMAF) and m900-MMAE. M860-MMAF is a sugar keto conjugation ADC targeting the HER2 receptor [128], and m909-ADC is a thiol conjugation targeting the folate receptor β [129] (Figure 6A). Their APRs were predicted by Aggrescan, TANGO, WALTZ and Amylpred2, and the hydrophobicity was empirically calculated by measuring the non-polar surface area of both antibodies and drug molecules [130]. Results showed that linker payload contributed to APRs, as well as the overall hydrophobicity (Figure 6B). Due to the increased hydrophobicity, ADCs are more prone to aggregation compared to the parent mAbs. Guo et al. have reported that an ADC with a maleimidocaproyl linker and an auristatin payload are more prone to aggregate during thermal stress than the parent mAb, although they have similar secondary and tertiary structures [131]. Similar results were obtained by our group for m860-MMAF (Figure 6C). To reduce the hydrophobicity of ADC, some groups have studied replacing the hydrophobic linkers with hydrophilic linkers or PEGylation linkers. Zhao et al. have incorporated sulfonate- or PEG-containing hydrophilic linkers into antibody maytansinoid conjugates to achieve high DAR without aggregation and low non-specificity [132]. Lyon et al. have exploited a hydrophilic glucuronide linker in PEGylated ADCs to decrease the hydrophobicity of ADC and to extend its in vivo half-life [133].

The one-pot chemical conjugation of mAb and drug reactants usually produces heterogeneous ADC products. In the case of thiol-maleimide chemistry, the extent to which the inter-chain disulfide bonds were reduced determined the quantity of drugs attaching to mAb, which led to the ADC species with different DARs ranging from 0 to 8. DAR constitutes a major concern for designing ADC therapeutics, since DAR tightly impacts the aggregation propensity, in vivo potency and serum stability of ADC. ADC with high DAR usually has higher aggregation propensity due to the increased hydrophobicity conferred by the hydrophobic drug molecules. Guo et al. have found that their ADC with DAR6 species may exist in a multimeric state, while DAR2 and DAR4 species likely exist in monomeric forms under ambient conditions [127]. Beckley et al. have reported that their ADC

aggregate mainly contained the high DAR species of 6–8 [124]. Meanwhile, ADC with high DAR was reported to be subjected to more structural perturbations, causing the destabilization of mAb, as exampled by Adem et al. showing that high DAR species readily experienced aggregation and fragmentation under stress conditions, such as high ionic strength buffer, due to the fewer inter-chain disulfide bonds [134]. The destabilized structure of high DAR species could probably reconcile its fast plasma clearance [133] when encountered with protease degradation. Besides, Pan et al. have reported that ADC bears distinct conformation at the C_H2–C_H3 interface compared to the parent antibody [135], which may disturb its interaction with FcRn and impact the serum half-life of ADC. In addition, the authors also found that the low hinge region of the C_H2 domain became more solvent exposed in ADC than in the naked mAb by using hydrogen/deuterium exchange mass spectrometry (HDX-MS). The conformation alteration of C_H2 domain was also confirmed by Beckley et al., showing that the C_H2 domains in ADC with high DAR bear low stability and rapidly form aggregates at 40 °C [124]. The fact that the higher DAR could pose more structural alterations is further supported by our molecular docking (PatchDock simulations [136]) studies of m860-MMAF and m900-MMAE with different DARs. Our results showed that the propensity of ADCs associating with each other (docking score) positively correlated with the DAR (Figure 6D), indicating that the higher DAR species presents a conformation that is more prone to associate with each other.

Figure 6. The aggregation of ADC. (**A**) The chemical structures of monomethyl auristatin F (MMAF) and E (MMAE) and their respective linkers and spacers; (**B**) The molecular surface diagrams of antibody drug conjugates (ADCs) with the predicted aggregation-prone regions (APRs) colored in red. Homology models of the ADCs M860-MMAE and M909-MMAF have the maximum number of four drugs for M860-MMAF and eight drugs for M909-MMAE. The APRs of the antibody are the consensus residues predicted by four combinations of calculation programs, Aggrescan, TANGO, WALTZ and Amylpred2. The APRs of the drug molecules are the nonpolar carbon atoms; (**C**) The aggregation tendencies of M860 mAb and M860-ADC measured using the dynamic light scattering test to determine the size distribution profile after seven days of incubation at 4 °C and 37 °C (1 mg/mL of PBS buffer). (**D**) The docking scores of M860-MMAF and M909-MMAE with drug-antibody ratios (DARs). The average docking scores of M860-MMAF and M909-MMAE were simulated and calculated using PatchDock.

Collectively, the aggregation of ADC is a function of the naked antibody, linker payload, conjugation chemistry and DAR, as well as the formulation solution, all of which need to be considered when trying to increase the aggregation resistance of ADCs.

5. Conclusions and Outlooks

The aggregation of antibody domains has been extensively studied in order to expedite the screening of aggregation-resistant IgG in the early development process. However, caution is needed since aggregation resistance strategies for antibody domains are not always successfully transferred onto full-sized IgG. Usually, the strategies for improving the stability of the constant domain could be applied to full-length IgG. For example, our group has found that the stabilization mutations at the C_H1/C_κ heterodimer interface (S64E/S66V of C_H1 and S69L/T71S of C_κ) identified from the Fab-like format could be applied onto the full-length IgG-like format (4Dm2m) [113]. Kelly et al. have found the mutations of Q295F/Y296A stabilizing the C_H2 domain through interaction with glycan could also confer full-sized IgG decreased low pH-induced aggregation [117]. By contrast, transferring of aggregation properties from variable domains to full-sized IgGs is often complicated by the antigen binding and domain-domain associations. Pepinsky et al. have reported that the anti-LINGO-1 Mab Li33 selected from the Fab phage library had poor solubility when converted into IgG1 format [67]. Another example came from Daniel Christ's group, showing that the aggregation resistance benefit derived from mutating residues in CDR1 of V_H and CDR2 of V_L into negatively-charged residues, such as Asp and Glu, could successfully be transferred into scFv format [63]. The scFv trastuzumab with double mutations of 30D/52D not only resisted aggregation, but also retained high binding affinity to HER2. Furthermore, these mutations have been incorporated into the full-length trastuzumab IgG1 without disturbing the antigen binding and the biological functions. However, it remains to be seen whether these mutations alleviating aggregation of the single domain antibody could also improve the solubility of the full-length IgG1.

One should also keep in mind that the same forces promoting protein aggregation also operate in protein folding and interactions. As such, one needs to balance between decreasing aggregation and maintaining the correct folding and specific binding when rationally-designing aggregation-resistant antibodies. In our development of aggregation-resistant m36.4 variants, we have found that although some mutants indeed gained decreased aggregation according to the dynamic light scattering (DLS) profiles, their unfolding T_ms was significantly compromised as measured by the temperature ramping CD spectra. Meanwhile, mitigating aggregation through engineering the IgG sequence and structure by mutagenesis bears risks to generate new B- and T-cell epitopes and to cause immunogenicity. While the mutations occurring at CDRs could induce an anti-idiotype response and lead to the neutralization of the therapeutics, mutations at the constant domain could evoke the "binding antibodies" response in the host and compromise the efficacy of therapeutics due to the changed pharmacokinetics [137]. Thus, the risks of eliciting immunogenicity should be carefully considered when designing aggregation-resistant mutations, although many CDR mutations are reported to cause little or no immunogenicity in clinical trials [138]. An effective approach to limit immunogenicity is germlining to the human counterpart. Thus, the mutations of IgG sequences should be as few as possible. Fortunately, the potential of introducing T-cell epitopes by mutations could now be predicted in silico due to the mapping of their interactions with the human leukocyte antigen (HLA) class II molecules [139]. Besides, one should pursue exploiting the benefit of the hydrophilicity of the carbohydrate and PEG (polyethylene glycol) to combat the aggregation propensity of IgG, since glycosylation and PEGylation of IgG have a much decreased possibility to induce immunogenicity [140,141].

The antibody multi-domain nature defines its aggregation process as pathway dependent, which is a function of a combination of IgG sequences and structures, as well as the solution conditions. While most designs focus on improving thermostability and alleviating the aggregation of antibody fragments in a lumped assay, such as turbidity and dye binding, few studies are

dedicated to clarifying to what extent the engineering strategies impact the formation of the aggregate intermediates or the soluble irreversible aggregate species. Similarly, few existing computational programs account for the aggregation mechanism of the antibody although they could predict APRs in antibody sequences. Furthermore, calculation tools have failed to incorporate the external solution condition into the aggregation prediction. Although several programs have indeed considered the physiological conditions in the prediction, therapeutic antibodies are not usually produced, stored and administered under physiological conditions. Thus, these calculation methods poorly learn the antibody aggregation-resistant engineering in some cases. The studies of soluble aggregate intermediates, aggregation kinetics and mechanisms, as well as their dependence on the solution conditions would guide the antibody formulation optimization, which should be extensively pursued in future research work.

In summary, the aggregation propensities and aggregation-resistance strategies for antibody domains have been extensively investigated. Future studies are worth understanding the aggregation mechanisms for full-sized antibodies and their dependence on the environment, which could expedite antibody therapeutics development.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Creighton, T.E. *Protein Folding*; WH Freeman and Company: New York, NY, USA, 1992.
2. Lindorff-Larsen, K.; Piana, S.; Dror, R.O.; Shaw, D.E. How fast-folding proteins fold. *Science* **2011**, *334*, 517–520. [CrossRef] [PubMed]
3. Dobson, C.M.; Karplus, M. The fundamentals of protein folding: Bringing together theory and experiment. *Curr. Opin. Struct. Biol.* **1999**, *9*, 92–101. [CrossRef]
4. James, L.C.; Tawfik, D.S. Conformational diversity and protein evolution—A 60-year-old hypothesis revisited. *Trends Biochem. Sci.* **2003**, *28*, 361–368. [CrossRef]
5. Chiti, F.; Dobson, C.M. Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **2009**, *5*, 15–22. [CrossRef] [PubMed]
6. Dobson, C.M. Protein folding and misfolding. *Nature* **2003**, *426*, 884–890. [CrossRef] [PubMed]
7. Roberts, C.J. Therapeutic protein aggregation: Mechanisms, design, and control. *Trends Biotechnol.* **2014**, *32*, 372–380. [CrossRef] [PubMed]
8. Dobson, C.M. The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc. Lond. Ser. B* **2001**, *356*, 133–145. [CrossRef] [PubMed]
9. Chiti, F.; Taddei, N.; Baroni, F.; Capanni, C.; Stefani, M.; Ramponi, G.; Dobson, C.M. Kinetic partitioning of protein folding and aggregation. *Nat. Struct. Biol.* **2002**, *9*, 137–143. [CrossRef] [PubMed]
10. Reichert, J.M. Monoclonal antibodies as innovative therapeutics. *Curr. Pharm. Biotechnol.* **2008**, *9*, 423–430. [CrossRef] [PubMed]
11. Reichert, J.M.; Rosensweig, C.J.; Faden, L.B.; Dewitz, M.C. Monoclonal antibody successes in the clinic. *Nat. Biotechnol.* **2005**, *23*, 1073–1078. [CrossRef] [PubMed]
12. Lowe, D.; Dudgeon, K.; Rouet, R.; Schofield, P.; Jermutus, L.; Christ, D. Aggregation, stability, and formulation of human antibody therapeutics. *Adv. Protein Chem. Struct. Biol.* **2011**, *84*, 41–61. [PubMed]
13. Perchiacca, J.M.; Tessier, P.M. Engineering aggregation-resistant antibodies. *Annu. Rev. Chem. Biomol. Eng.* **2012**, *3*, 263–286. [CrossRef] [PubMed]
14. Nishi, H.; Miyajima, M.; Nakagami, H.; Noda, M.; Uchiyama, S.; Fukui, K. Phase separation of an IgG1 antibody solution under a low ionic strength condition. *Pharm. Res.* **2010**, *27*, 1348–1360. [CrossRef] [PubMed]

15. Liu, J.; Nguyen, M.D.; Andya, J.D.; Shire, S.J. Reversible self-association increases the viscosity of a concentrated monoclonal antibody in aqueous solution. *J. Pharm. Sci.* **2005**, *94*, 1928–1940. [CrossRef] [PubMed]

16. Richardson, J.S.; Richardson, D.C. Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2754–2759. [CrossRef] [PubMed]

17. Wu, S.J.; Luo, J.; O'Neil, K.T.; Kang, J.; Lacy, E.R.; Canziani, G.; Baker, A.; Huang, M.; Tang, Q.M.; Raju, T.S.; et al. Structure-based engineering of a monoclonal antibody for improved solubility. *Protein Eng. Des. Sel.* **2010**, *23*, 643–651. [CrossRef] [PubMed]

18. Wang, X.; Das, T.K.; Singh, S.K.; Kumar, S. Potential aggregation prone regions in biotherapeutics: A survey of commercial monoclonal antibodies. *mAbs* **2009**, *1*, 254–267. [CrossRef] [PubMed]

19. Nishi, H.; Miyajima, M.; Wakiyama, N.; Kubota, K.; Hasegawa, J.; Uchiyama, S.; Fukui, K. Fc domain mediated self-association of an IgG1 monoclonal antibody under a low ionic strength condition. *J. Biosci. Bioeng.* **2011**, *112*, 326–332. [CrossRef] [PubMed]

20. Nezlin, R. Interactions between immunoglobulin G molecules. *Immunol. Lett.* **2010**, *132*, 1–5. [CrossRef] [PubMed]

21. Shire, S.J. Formulation and manufacturability of biologics. *Curr. Opin. Biotechnol.* **2009**, *20*, 708–714. [CrossRef] [PubMed]

22. Daugherty, A.L.; Mrsny, R.J. Formulation and delivery issues for monoclonal antibody therapeutics. *Adv. Drug Deliv. Rev.* **2006**, *58*, 686–706. [CrossRef] [PubMed]

23. De Groot, A.S.; Scott, D.W. Immunogenicity of protein therapeutics. *Trends Immunol.* **2007**, *28*, 482–490. [CrossRef] [PubMed]

24. Singh, S.K. Impact of product-related factors on immunogenicity of biotherapeutics. *J. Pharm. Sci.* **2011**, *100*, 354–387. [CrossRef] [PubMed]

25. Hotzel, I.; Theil, F.P.; Bernstein, L.J.; Prabhu, S.; Deng, R.; Quintana, L.; Lutman, J.; Sibia, R.; Chan, P.; Bumbaca, D.; et al. A strategy for risk mitigation of antibodies with fast clearance. *mAbs* **2012**, *4*, 753–760. [CrossRef] [PubMed]

26. Tsumoto, K.; Ejima, D.; Kita, Y.; Arakawa, T. Review: Why is arginine effective in suppressing aggregation? *Protein Pept. Lett.* **2005**, *12*, 613–619. [CrossRef] [PubMed]

27. Manning, M.C.; Matsuura, J.E.; Kendrick, B.S.; Meyer, J.D.; Dormish, J.J.; Vrkljan, M.; Ruth, J.R.; Carpenter, J.F.; Sheftert, E. Approaches for increasing the solution stability of proteins. *Biotechnol. Bioeng.* **1995**, *48*, 506–512. [CrossRef] [PubMed]

28. Roberts, C.J. Non-native protein aggregation kinetics. *Biotechnol. Bioeng.* **2007**, *98*, 927–938. [CrossRef] [PubMed]

29. Wu, H.; Kroe-Barrett, R.; Singh, S.; Robinson, A.S.; Roberts, C.J. Competing aggregation pathways for monoclonal antibodies. *FEBS Lett.* **2014**, *588*, 936–941. [CrossRef] [PubMed]

30. Ramirez-Alvarado, M.; Merkel, J.S.; Regan, L. A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8979–8984. [CrossRef] [PubMed]

31. Vermeer, A.W.; Norde, W. The thermal stability of immunoglobulin: Unfolding and aggregation of a multi-domain protein. *Biophys. J.* **2000**, *78*, 394–404. [CrossRef]

32. Fink, A.L. Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **1998**, *3*, R9–R23. [CrossRef]

33. Vendruscolo, M.; Paci, E. Protein folding: Bringing theory and experiment closer together. *Curr. Opin. Struct. Biol.* **2003**, *13*, 82–87. [CrossRef]

34. De Simone, A.; Dhulesia, A.; Soldi, G.; Vendruscolo, M.; Hsu, S.T.; Chiti, F.; Dobson, C.M. Experimental free energy surfaces reveal the mechanisms of maintenance of protein solubility. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 21057–21062. [CrossRef] [PubMed]

35. Sahin, E.; Grillo, A.O.; Perkins, M.D.; Roberts, C.J. Comparative effects of pH and ionic strength on protein-protein interactions, unfolding, and aggregation for IgG1 antibodies. *J. Pharm. Sci.* **2010**, *99*, 4830–4848. [CrossRef] [PubMed]

36. Wang, W. Protein aggregation and its inhibition in biopharmaceutics. *Int. J. Pharm.* **2005**, *289*, 1–30. [CrossRef] [PubMed]

37. Russel, W.B.; Saville, D.A.; Schowalter, W.R. *Colloidal Dispersions*; Cambridge University Press: Cambridge, UK, 1992.

38. Geng, S.B.; Cheung, J.K.; Narasimhan, C.; Shameem, M.; Tessier, P.M. Improving monoclonal antibody selection and engineering using measurements of colloidal protein interactions. *J. Pharm. Sci.* **2014**, *103*, 3356–3363. [CrossRef] [PubMed]

39. Roberts, C.J. Kinetics of Irreversible Protein Aggregation: Analysis of Extended Lumry-Eyring Models and Implications for Predicting Protein Shelf Life. *J. Phys. Chem. B* **2003**, *107*, 1194–1207. [CrossRef]

40. Lee, C.C.; Perchiacca, J.M.; Tessier, P.M. Toward aggregation-resistant antibodies by design. *Trends Biotechnol.* **2013**, *31*, 612–620. [CrossRef] [PubMed]

41. Booth, D.R.; Sunde, M.; Bellotti, V.; Robinson, C.V.; Hutchinson, W.L.; Fraser, P.E.; Hawkins, P.N.; Dobson, C.M.; Radford, S.E.; Blake, C.C.; et al. Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* **1997**, *385*, 787–793. [CrossRef] [PubMed]

42. Rosen, D.R.; Siddique, T.; Patterson, D.; Figlewicz, D.A.; Sapp, P.; Hentati, A.; Donaldson, D.; Goto, J.; O'Regan, J.P.; Deng, H.X.; et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **1993**, *362*, 59–62. [CrossRef] [PubMed]

43. Broome, B.M.; Hecht, M.H. Nature disfavors sequences of alternating polar and non-polar amino acids: Implications for amyloidogenesis. *J. Mol. Biol.* **2000**, *296*, 961–968. [CrossRef] [PubMed]

44. Schwartz, R.; Istrail, S.; King, J. Frequencies of amino acid strings in globular protein sequences indicate suppression of blocks of consecutive hydrophobic residues. *Protein Sci.: Publ. Protein Soc.* **2001**, *10*, 1023–1031. [CrossRef] [PubMed]

45. Cellmer, T.; Bratko, D.; Prausnitz, J.M.; Blanch, H.W. Protein aggregation in silico. *Trends Biotechnol.* **2007**, *25*, 254–261. [CrossRef] [PubMed]

46. Gupta, P.; Hall, C.K.; Voegler, A.C. Effect of denaturant and protein concentrations upon protein refolding and aggregation: A simple lattice model. *Protein Sci. Publ. Protein Soc.* **1998**, *7*, 2642–2652. [CrossRef] [PubMed]

47. Ma, B.; Nussinov, R. Molecular dynamics simulations of alanine rich beta-sheet oligomers: Insight into amyloid formation. *Protein Sci. Publ. Protein Soc.* **2002**, *11*, 2335–2350. [CrossRef] [PubMed]

48. Wolf, M.G.; Jongejan, J.A.; Laman, J.D.; de Leeuw, S.W. Quantitative prediction of amyloid fibril growth of short peptides from simulations: Calculating association constants to dissect side chain importance. *J. Am. Chem. Soc.* **2008**, *130*, 15772–15773. [CrossRef] [PubMed]

49. De Groot, N.S.; Pallares, I.; Aviles, F.X.; Vendrell, J.; Ventura, S. Prediction of "hot spots" of aggregation in disease-linked polypeptides. *BMC Struct. Biol.* **2005**, *5*. [CrossRef] [PubMed]

50. Monsellier, E.; Ramazzotti, M.; de Laureto, P.P.; Tartaglia, G.G.; Taddei, N.; Fontana, A.; Vendruscolo, M.; Chiti, F. The distribution of residues in a polypeptide sequence is a determinant of aggregation optimized by evolution. *Biophys. J.* **2007**, *93*, 4382–4391. [CrossRef] [PubMed]

51. Zhang, Z.; Chen, H.; Lai, L. Identification of amyloid fibril-forming segments based on structure and residue-based statistical potential. *Bioinformatics* **2007**, *23*, 2218–2225. [CrossRef] [PubMed]

52. Agrawal, N.J.; Kumar, S.; Wang, X.; Helk, B.; Singh, S.K.; Trout, B.L. Aggregation in protein-based biotherapeutics: Computational studies and tools to identify aggregation-prone regions. *J. Pharm. Sci.* **2011**, *100*, 5081–5095. [CrossRef] [PubMed]

53. Chennamsetty, N.; Helk, B.; Voynov, V.; Kayser, V.; Trout, B.L. Aggregation-prone motifs in human immunoglobulin G. *J. Mol. Biol.* **2009**, *391*, 404–413. [CrossRef] [PubMed]

54. Chennamsetty, N.; Voynov, V.; Kayser, V.; Helk, B.; Trout, B.L. Design of therapeutic proteins with enhanced stability. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11937–11942. [CrossRef] [PubMed]

55. Schroeder, H.W., Jr.; Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **2010**, *125*, S41–S52. [CrossRef] [PubMed]

56. Alamyar, E.; Giudicelli, V.; Duroux, P.; Lefranc, M.P. Antibody V and C domain sequence, structure, and interaction analysis with special reference to IMGT(R). *Methods Mol. Biol.* **2014**, *1131*, 337–381. [PubMed]

57. Ionescu, R.M.; Vlasak, J.; Price, C.; Kirchmeier, M. Contribution of variable domains to the stability of humanized IgG1 monoclonal antibodies. *J. Pharm. Sci.* **2008**, *97*, 1414–1426. [CrossRef] [PubMed]

58. Latypov, R.F.; Hogan, S.; Lau, H.; Gadgil, H.; Liu, D. Elucidation of acid-induced unfolding and aggregation of human immunoglobulin IgG1 and IgG2 Fc. *J. Biol. Chem.* **2012**, *287*, 1381–1396. [CrossRef] [PubMed]

59. Weiss, W.F.; Young, T.M.; Roberts, C.J. Principles, approaches, and challenges for predicting protein aggregation rates and shelf life. *J. Pharm. Sci.* **2009**, *98*, 1246–1277. [CrossRef] [PubMed]

60. Telikepalli, S.N.; Kumru, O.S.; Kalonia, C.; Esfandiary, R.; Joshi, S.B.; Middaugh, C.R.; Volkin, D.B. Structural characterization of IgG1 mAb aggregates and particles generated under various stress conditions. *J. Pharm. Sci.* **2014**, *103*, 796–809. [CrossRef] [PubMed]

61. Trivedi, M.V.; Laurence, J.S.; Siahaan, T.J. The role of thiols and disulfides on protein stability. *Curr. Protein Pept. Sci.* **2009**, *10*, 614–625. [CrossRef] [PubMed]

62. Buchanan, A.; Clementel, V.; Woods, R.; Harn, N.; Bowen, M.A.; Mo, W.; Popovic, B.; Bishop, S.M.; Dall'Acqua, W.; Minter, R.; et al. Engineering a therapeutic IgG molecule to address cysteinylation, aggregation and enhance thermal stability and expression. *mAbs* **2013**, *5*, 255–262. [CrossRef] [PubMed]

63. Dudgeon, K.; Rouet, R.; Kokmeijer, I.; Schofield, P.; Stolp, J.; Langley, D.; Stock, D.; Christ, D. General strategy for the generation of human antibody variable domains with increased aggregation resistance. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10879–10884. [CrossRef] [PubMed]

64. Kayser, V.; Chennamsetty, N.; Voynov, V.; Forrer, K.; Helk, B.; Trout, B.L. Glycosylation influences on the aggregation propensity of therapeutic monoclonal antibodies. *Biotechnol. J.* **2011**, *6*, 38–44. [CrossRef] [PubMed]

65. Krapp, S.; Mimura, Y.; Jefferis, R.; Huber, R.; Sondermann, P. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J. Mol. Biol.* **2003**, *325*, 979–989. [CrossRef]

66. Schaefer, J.V.; Plückthun, A. Engineering Aggregation Resistance in IgG by Two Independent Mechanisms: Lessons from Comparison of Pichia pastoris and Mammalian Cell Expression. *J. Mol. Biol.* **2012**, *417*, 309–335. [CrossRef] [PubMed]

67. Pepinsky, R.B.; Silvian, L.; Berkowitz, S.A.; Farrington, G.; Lugovskoy, A.; Walus, L.; Eldredge, J.; Capili, A.; Mi, S.; Graff, C.; et al. Improving the solubility of anti-LINGO-1 monoclonal antibody Li33 by isotype switching and targeted mutagenesis. *Protein Sci.: Publ. Protein Soc.* **2010**, *19*, 954–966. [CrossRef] [PubMed]

68. Nezlin, R.; Ghetie, V. Interactions of immunoglobulins outside the antigen-combining site. *Adv. Immunol.* **2004**, *82*, 155–215. [PubMed]

69. Nezlin, R. Internal movements in immunoglobulin molecules. *Adv. Immunol.* **1990**, *48*, 1–40. [PubMed]

70. Tiller, K.E.; Tessier, P.M. Advances in Antibody Design. *Annu. Rev. Biomed. Eng.* **2015**, *17*, 191–216. [CrossRef] [PubMed]

71. Rouet, R.; Lowe, D.; Christ, D. Stability engineering of the human antibody repertoire. *FEBS Lett.* **2014**, *588*, 269–277. [CrossRef] [PubMed]

72. Wang, W.; Singh, S.; Zeng, D.L.; King, K.; Nema, S. Antibody structure, instability, and formulation. *J. Pharm. Sci.* **2007**, *96*, 1–26. [CrossRef] [PubMed]

73. Tischenko, V.M.; Abramov, V.M.; Zav'yalov, V.P. Investigation of the cooperative structure of Fc fragments from myeloma immunoglobulin G. *Biochemistry* **1998**, *37*, 5576–5581. [CrossRef] [PubMed]

74. Goto, Y.; Ichimura, N.; Hamaguchi, K. Effects of ammonium sulfate on the unfolding and refolding of the variable and constant fragments of an immunoglobulin light chain. *Biochemistry* **1988**, *27*, 1670–1677. [CrossRef] [PubMed]

75. Zav'yalov, V.P.; Tishchenko, V.M. Mechanisms of generation of antibody diversity as a cause for natural selection of homoiothermal animals in the process of evolution. *Scand. J. Immunol.* **1991**, *33*, 755–762. [CrossRef] [PubMed]

76. Chen, W.; Gong, R.; Ying, T.; Prabakaran, P.; Zhu, Z.; Feng, Y.; Dimitrov, D.S. Discovery of novel candidate therapeutics and diagnostics based on engineered human antibody domains. *Curr. Drug Discov. Technol.* **2014**, *11*, 28–40. [CrossRef] [PubMed]

77. Hussack, G.; Keklikian, A.; Alsughayyir, J.; Hanifi-Moghaddam, P.; Arbabi-Ghahroudi, M.; van Faassen, H.; Hou, S.T.; Sad, S.; MacKenzie, R.; Tanha, J. A VL single-domain antibody library shows a high-propensity to yield non-aggregating binders. *Protein Eng. Des. Sel.* **2012**, *25*, 313–318. [CrossRef] [PubMed]

78. Goto, Y.; Hamaguchi, K. The role of the intrachain disulfide bond in the conformation and stability of the constant fragment of the immunoglobulin light chain. *J. Biochem.* **1979**, *86*, 1433–1441. [PubMed]

79. Chothia, C.; Janin, J. Principles of protein-protein recognition. *Nature* **1975**, *256*, 705–708. [CrossRef] [PubMed]

80. Arbabi-Ghahroudi, M.; To, R.; Gaudette, N.; Hirama, T.; Ding, W.; MacKenzie, R.; Tanha, J. Aggregation-resistant VHs selected by in vitro evolution tend to have disulfide-bonded loops and acidic isoelectric points. *Protein Eng. Des. Sel.* **2009**, *22*, 59–66. [CrossRef] [PubMed]

81. Perchiacca, J.M.; Ladiwala, A.R.A.; Bhattacharya, M.; Tessier, P.M. Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions. *Protein Eng. Des. Sel.* **2012**, *25*, 591–602. [CrossRef] [PubMed]

82. Perchiacca, J.M.; Bhattacharya, M.; Tessier, P.M. Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. *Proteins* **2011**, *79*, 2637–2647. [CrossRef] [PubMed]

83. Greenberg, A.S.; Avila, D.; Hughes, M.; Hughes, A.; McKinney, E.C.; Flajnik, M.F. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature* **1995**, *374*, 168–173. [CrossRef] [PubMed]

84. Davies, J.; Riechmann, L. Single antibody domains as small recognition units: Design and in vitro antigen selection of camelized, human VH domains with improved protein stability. *Protein Eng.* **1996**, *9*, 531–537. [CrossRef] [PubMed]

85. Kim, D.Y.; Kandalaft, H.; Ding, W.; Ryan, S.; van Faassen, H.; Hirama, T.; Foote, S.J.; MacKenzie, R.; Tanha, J. Disulfide linkage engineering for improving biophysical properties of human VH domains. *Protein Eng. Des. Sel.* **2012**, *25*, 581–590. [CrossRef] [PubMed]

86. Kim, D.Y.; To, R.; Kandalaft, H.; Ding, W.; van Faassen, H.; Luo, Y.; Schrag, J.D.; St-Amant, N.; Hefford, M.; Hirama, T.; et al. Antibody light chain variable domains and their biophysically improved versions for human immunotherapy. *mAbs* **2014**, *6*, 219–235. [CrossRef] [PubMed]

87. Barthelemy, P.A.; Raab, H.; Appleton, B.A.; Bond, C.J.; Wu, P.; Wiesmann, C.; Sidhu, S.S. Comprehensive Analysis of the Factors Contributing to the Stability and Solubility of Autonomous Human VH Domains. *J. Biol. Chem.* **2008**, *283*, 3639–3654. [CrossRef] [PubMed]

88. Davies, J.; Riechmann, L. 'Camelising' human antibody fragments: NMR studies on VH domains. *FEBS Lett.* **1994**, *339*, 285–290. [CrossRef]

89. Langedijk, A.C.; Honegger, A.; Maat, J.; Planta, R.J.; van Schaik, R.C.; Pluckthun, A. The nature of antibody heavy chain residue H6 strongly influences the stability of a VH domain lacking the disulfide bridge. *J. Mol. Biol.* **1998**, *283*, 95–110. [CrossRef] [PubMed]

90. Gorlani, A.; Hulsik, D.L.; Adams, H.; Vriend, G.; Hermans, P.; Verrips, T. Antibody engineering reveals the important role of J segments in the production efficiency of llama single-domain antibodies in Saccharomyces cerevisiae. *Protein Eng. Des. Sel.* **2012**, *25*, 39–46. [CrossRef] [PubMed]

91. Spada, S.; Honegger, A.; Pluckthun, A. Reproducing the natural evolution of protein structural features with the selectively infective phage (SIP) technology. The kink in the first strand of antibody kappa domains. *J. Mol. Biol.* **1998**, *283*, 395–407. [CrossRef] [PubMed]

92. Ewert, S.; Huber, T.; Honegger, A.; Pluckthun, A. Biophysical properties of human antibody variable domains. *J. Mol. Biol.* **2003**, *325*, 531–553. [CrossRef]

93. Nilson, B.H.; Solomon, A.; Bjorck, L.; Akerstrom, B. Protein L from Peptostreptococcus magnus binds to the kappa light chain variable domain. *J. Biol. Chem.* **1992**, *267*, 2234–2239. [PubMed]

94. Steipe, B.; Schiller, B.; Pluckthun, A.; Steinbacher, S. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* **1994**, *240*, 188–192. [CrossRef] [PubMed]

95. Ignatovich, O.; Jespers, L.; Tomlinson, I.M.; de Wildt, R.M.T. Creation of the Large and Highly Functional Synthetic Repertoire of Human VH and Vκ Domain Antibodies. In *Single Domain Antibodies: Methods and Protocols*; Saerens, D., Muyldermans, S., Eds.; Humana Press: Totowa, NJ, USA, 2012; pp. 39–63.

96. Chen, W.; Zhu, Z.; Feng, Y.; Xiao, X.; Dimitrov, D.S. Construction of a Large Phage-Displayed Human Antibody Domain Library with a Scaffold Based On a Newly Identified Highly Soluble, Stable Heavy Chain Variable Domain. *J. Mol. Biol.* **2008**, *382*, 779–789. [CrossRef] [PubMed]

97. Colby, D.W.; Garg, P.; Holden, T.; Chao, G.; Webster, J.M.; Messer, A.; Ingram, V.M.; Wittrup, K.D. Development of a Human Light Chain Variable Domain (VL) Intracellular Antibody Specific for the Amino Terminus of Huntingtin via Yeast Surface Display. *J. Mol. Biol.* **2004**, *342*, 901–912. [CrossRef] [PubMed]

98. Gil, D.; Schrum, A.G. Strategies to stabilize compact folding and minimize aggregation of antibody-based fragments. *Adv. Biosci. Biotechnol.* **2013**, *4*, 73–84. [CrossRef] [PubMed]

99. Worn, A.; Pluckthun, A. Stability engineering of antibody single-chain Fv fragments. *J. Mol. Biol.* **2001**, *305*, 989–1010. [CrossRef] [PubMed]

100. Pluckthun, A. Mono- and bivalent antibody fragments produced in *Escherichia coli*: Engineering, folding and antigen binding. *Immunol. Rev.* **1992**, *130*, 151–188. [CrossRef] [PubMed]

101. Glockshuber, R.; Malia, M.; Pfitzinger, I.; Pluckthun, A. A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochemistry* **1990**, *29*, 1362–1367. [CrossRef] [PubMed]

102. Tan, P.H.; Sandmaier, B.M.; Stayton, P.S. Contributions of a highly conserved VH/VL hydrogen bonding interaction to scFv folding stability and refolding efficiency. *Biophys. J.* **1998**, *75*, 1473–1482. [CrossRef]

103. Zhu, Z.; Presta, L.G.; Zapata, G.; Carter, P. Remodeling domain interfaces to enhance heterodimer formation. *Protein Sci. Publ. Protein Soc.* **1997**, *6*, 781–788. [CrossRef] [PubMed]

104. Hakansson, M.; Linse, S. Protein reconstitution and 3D domain swapping. *Curr. Protein Pept. Sci.* **2002**, *3*, 629–642. [CrossRef] [PubMed]

105. Robinson, C.R.; Sauer, R.T. Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5929–5934. [CrossRef] [PubMed]

106. Arndt, K.M.; Muller, K.M.; Pluckthun, A. Factors influencing the dimer to monomer transition of an antibody single-chain Fv fragment. *Biochemistry* **1998**, *37*, 12918–12926. [CrossRef] [PubMed]

107. Nelson, A.L.; Reichert, J.M. Development trends for therapeutic antibody fragments. *Nat. Biotechnol.* **2009**, *27*, 331–337. [CrossRef] [PubMed]

108. Wang, T.; Kumru, O.S.; Yi, L.; Wang, Y.J.; Zhang, J.; Kim, J.H.; Joshi, S.B.; Middaugh, C.R.; Volkin, D.B. Effect of ionic strength and pH on the physical and chemical stability of a monoclonal antibody antigen-binding fragment. *J. Pharm. Sci.* **2013**, *102*, 2520–2537. [CrossRef] [PubMed]

109. Roux, K.H.; Tankersley, D.L. A view of the human idiotypic repertoire. Electron microscopic and immunologic analyses of spontaneous idiotype-anti-idiotype dimers in pooled human IgG. *J. Immunol.* **1990**, *144*, 1387–1395. [PubMed]

110. Kanai, S.; Liu, J.; Patapoff, T.W.; Shire, S.J. Reversible self-association of a concentrated monoclonal antibody solution mediated by Fab-Fab interaction that impacts solution viscosity. *J. Pharm. Sci.* **2008**, *97*, 4219–4227. [CrossRef] [PubMed]

111. Peters, S.J.; Smales, C.M.; Henry, A.J.; Stephens, P.E.; West, S.; Humphreys, D.P. Engineering an Improved IgG4 Molecule with Reduced Disulfide Bond Heterogeneity and Increased Fab Domain Thermal Stability. *J. Biol. Chem.* **2012**, *287*, 24525–24533. [CrossRef] [PubMed]

112. Teerinen, T.; Valjakka, J.; Rouvinen, J.; Takkinen, K. Structure-based stability engineering of the mouse IgG1 Fab fragment by modifying constant domains. *J. Mol. Biol.* **2006**, *361*, 687–697. [CrossRef] [PubMed]

113. Chen, W.; Bardhi, A.; Feng, Y.; Wang, Y.; Qi, Q.; Li, W.; Zhu, Z.; Dyba, M.A.; Ying, T.; Jiang, S.; et al. Improving the CH1-CK heterodimerization and pharmacokinetics of 4Dm2m, a novel potent CD4-antibody fusion protein against HIV-1. *mAbs* **2016**, *8*, 761–774. [CrossRef] [PubMed]

114. Gong, R.; Wang, Y.; Feng, Y.; Zhao, Q.; Dimitrov, D.S. Shortened engineered human antibody CH2 domains: Increased stability and binding to the human neonatal Fc receptor. *J. Biol. Chem.* **2011**, *286*, 27288–27293. [CrossRef] [PubMed]

115. DeLano, W.L.; Ultsch, M.H.; de Vos, A.M.; Wells, J.A. Convergent solutions to binding at a protein-protein interface. *Science* **2000**, *287*, 1279–1283. [CrossRef] [PubMed]

116. Kolenko, P.; Dohnálek, J.; Dušková, J.; Skálová, T.; Collard, R.; Hašek, J. New insights into intra- and intermolecular interactions of immunoglobulins: Crystal structure of mouse IgG2b-Fc at 2·1-Å resolution. *Immunology* **2009**, *126*, 378–385. [CrossRef] [PubMed]

117. Chen, W.; Kong, L.; Connelly, S.; Dendle, J.M.; Liu, Y.; Wilson, I.A.; Powers, E.T.; Kelly, J.W. Stabilizing the CH2 Domain of an Antibody by Engineering in an Enhanced Aromatic Sequon. *ACS Chem. Biol.* **2016**, *11*, 1852–1861. [CrossRef] [PubMed]

118. Fast, J.L.; Cordes, A.A.; Carpenter, J.F.; Randolph, T.W. Physical instability of a therapeutic Fc fusion protein: Domain contributions to conformational and colloidal stability. *Biochemistry* **2009**, *48*, 11724–11736. [CrossRef] [PubMed]

119. Gong, R.; Vu, B.K.; Feng, Y.; Prieto, D.A.; Dyba, M.A.; Walsh, J.D.; Prabakaran, P.; Veenstra, T.D.; Tarasov, S.G.; Ishima, R.; et al. Engineered human antibody constant domains with increased stability. *J. Biol. Chem.* **2009**, *284*, 14203–14210. [CrossRef] [PubMed]

120. Ying, T.; Chen, W.; Feng, Y.; Wang, Y.; Gong, R.; Dimitrov, D.S. Engineered soluble monomeric IgG1 CH3 domain: Generation, mechanisms of function, and implications for design of biological therapeutics. *J. Biol. Chem.* **2013**, *288*, 25154–25164. [CrossRef] [PubMed]

121. Wozniak-Knopp, G.; Stadlmann, J.; Ruker, F. Stabilisation of the Fc fragment of human IgG1 by engineered intradomain disulfide bonds. *PLoS ONE* **2012**, *7*. [CrossRef] [PubMed]

122. Scott, A.M.; Wolchok, J.D.; Old, L.J. Antibody therapy of cancer. *Nat. Rev. Cancer* **2012**, *12*, 278–287. [CrossRef] [PubMed]

123. Zolot, R.S.; Basu, S.; Million, R.P. Antibody-drug conjugates. *Nat. Rev. Drug Discov.* **2013**, *12*, 259–260. [CrossRef] [PubMed]

124. Beckley, N.S.; Lazzareschi, K.P.; Chih, H.-W.; Sharma, V.K.; Flores, H.L. Investigation into Temperature-Induced Aggregation of an Antibody Drug Conjugate. *Bioconjug. Chem.* **2013**, *24*, 1674–1683. [CrossRef] [PubMed]

125. Wakankar, A.A.; Feeney, M.B.; Rivera, J.; Chen, Y.; Kim, M.; Sharma, V.K.; Wang, Y.J. Physicochemical Stability of the Antibody-Drug Conjugate Trastuzumab-DM1: Changes due to Modification and Conjugation Processes. *Bioconjug. Chem.* **2010**, *21*, 1588–1595. [CrossRef] [PubMed]

126. Acchione, M.; Kwon, H.; Jochheim, C.M.; Atkins, W.M. Impact of linker and conjugation chemistry on antigen binding, Fc receptor binding and thermal stability of model antibody-drug conjugates. *mAbs* **2012**, *4*, 362–372. [CrossRef] [PubMed]

127. Guo, J.; Kumar, S.; Chipley, M.; Marcq, O.; Gupta, D.; Jin, Z.; Tomar, D.S.; Swabowski, C.; Smith, J.; Starkey, J.A.; et al. Characterization and Higher-Order Structure Assessment of an Interchain Cysteine-Based ADC: Impact of Drug Loading and Distribution on the Mechanism of Aggregation. *Bioconjug. Chem.* **2016**, *27*, 604–615. [CrossRef] [PubMed]

128. Zhu, Z.; Ramakrishnan, B.; Li, J.; Wang, Y.; Feng, Y.; Prabakaran, P.; Colantonio, S.; Dyba, M.A.; QasbaM, P.K.; Dimitrov, D.S. Site-specific antibody-drug conjugation through an engineered glycotransferase and a chemically reactive sugar. *mAbs* **2014**, *6*, 1190–1200. [CrossRef] [PubMed]

129. Feng, Y.; Shen, J.; Streaker, E.D.; Lockwood, M.; Zhu, Z.; Low, P.S.; Dimitrov, D.S. A folate receptor beta-specific human monoclonal antibody recognizes activated macrophage of rheumatoid patients and mediates antibody-dependent cell-mediated cytotoxicity. *Arthritis Res. Ther.* **2011**, *13*. [CrossRef] [PubMed]

130. Chothia, C. Hydrophobic bonding and accessible surface area in proteins. *Nature* **1974**, *248*, 338–339. [CrossRef] [PubMed]

131. Guo, J.; Kumar, S.; Prashad, A.; Starkey, J.; Singh, S.K. Assessment of Physical Stability of an Antibody Drug Conjugate by Higher Order Structure Analysis: Impact of Thiol- Maleimide Chemistry. *Pharm. Res.* **2014**, *31*, 1710–1723. [CrossRef] [PubMed]

132. Zhao, R.Y.; Wilhelm, S.D.; Audette, C.; Jones, G.; Leece, B.A.; Lazar, A.C.; Goldmacher, V.S.; Singh, R.; Kovtun, Y.; Widdison, W.C.; et al. Synthesis and Evaluation of Hydrophilic Linkers for Antibody-Maytansinoid Conjugates. *J. Med. Chemi.* **2011**, *54*, 3606–3623. [CrossRef] [PubMed]

133. Lyon, R.P.; Bovee, T.D.; Doronina, S.O.; Burke, P.J.; Hunter, J.H.; Neff-LaFord, H.D.; Jonas, M.; Anderson, M.E.; Setter, J.R.; Senter, P.D. Reducing hydrophobicity of homogeneous antibody-drug conjugates improves pharmacokinetics and therapeutic index. *Nat. Biotech.* **2015**, *33*, 733–735. [CrossRef] [PubMed]

134. Adem, Y.T.; Schwarz, K.A.; Duenas, E.; Patapoff, T.W.; Galush, W.J.; Esue, O. Auristatin Antibody Drug Conjugate Physical Instability and the Role of Drug Payload. *Bioconjug. Chem.* **2014**, *25*, 656–664. [CrossRef] [PubMed]

135. Pan, L.Y.; Salas-Solano, O.; Valliere-Douglass, J.F. Conformation and Dynamics of Interchain Cysteine-Linked Antibody-Drug Conjugates as Revealed by Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2014**, *86*, 2657–2664. [CrossRef] [PubMed]

136. Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H.J. PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33*, W363–W367. [CrossRef] [PubMed]

137. Harding, F.A.; Stickler, M.M.; Razo, J.; DuBridge, R.B. The immunogenicity of humanized and fully human antibodies: Residual immunogenicity resides in the CDR regions. *mAbs* **2010**, *2*, 256–265. [CrossRef] [PubMed]

138. Hwang, W.Y.; Foote, J. Immunogenicity of engineered antibodies. *Methods* **2005**, *36*, 3–10. [CrossRef] [PubMed]

139. Van Walle, I.; Gansemans, Y.; Parren, P.W.; Stas, P.; Lasters, I. Immunogenicity screening in protein drug development. *Expert Opin. Biol. Ther.* **2007**, *7*, 405–418. [CrossRef] [PubMed]

140. Sinclair, A.M.; Elliott, S. Glycoengineering: The effect of glycosylation on the properties of therapeutic proteins. *J. Pharm. Sci.* **2005**, *94*, 1626–1635. [CrossRef] [PubMed]

141. Veronese, F.M.; Mero, A. The impact of PEGylation on biological therapies. *BioDrugs Clin. Immunother. Biopharm. Gene Ther.* **2008**, *22*, 315–329. [CrossRef]